

控制与决策

Control and Decision

考虑多维修队合作策略的灾后路网修复调度与路由联合优化

谭晓燕, 毛新华, 周继彪, 徐婷

引用本文:

谭晓燕, 毛新华, 周继彪, 等. 考虑多维修队合作策略的灾后路网修复调度与路由联合优化[J]. *控制与决策*, 2024, 39(12): 4151-4159.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2023.1724>

您可能感兴趣的其他文章

Articles you may be interested in

一种面向严重受损路网的抢修队调度算法

An algorithm for repair crew scheduling on severely damaged road network

控制与决策. 2021, 36(7): 1663-1671 <https://doi.org/10.13195/j.kzyjc.2019.1582>

两部件系统视情维修与生产调度的联合优化模型

Joint optimization model for condition-based maintenance and production scheduling of two-component systems

控制与决策. 2021, 36(6): 1377-1386 <https://doi.org/10.13195/j.kzyjc.2019.1357>

基于随机系数回归模型的退化过程及维修策略

Degradation process and maintenance planning based on random coefficient regression model

控制与决策. 2021, 36(3): 754-760 <https://doi.org/10.13195/j.kzyjc.2019.1087>

基于马尔可夫过程的多部件系统劣化状态空间划分模型

Multi-component system state space partition model based on Markov process

控制与决策. 2021, 36(2): 418-428 <https://doi.org/10.13195/j.kzyjc.2019.0480>

基于生成对抗网络的大规模路网交通流预测算法

Traffic flow forecasting algorithm for large-scale road network based on GAN

控制与决策. 2021, 36(12): 2937-2945 <https://doi.org/10.13195/j.kzyjc.2020.0333>

考虑多维修队合作策略的灾后路网修复调度与路由联合优化

谭晓燕^{1,2,3}, 毛新华^{1,2,3†}, 周继彪⁴, 徐婷¹

- (1. 长安大学 运输工程学院, 西安 710064;
2. 陕西省交通基础设施建设与管理数字化工程研究中心, 西安 710064;
3. 西安市交通基础设施建设与管理数字化重点实验室, 西安 710064;
4. 宁波市高等级公路建设管理中心, 浙江 宁波 315199)

摘要: 灾后路网修复决策研究中“一个受损路段只能由一个维修队修复”的常用约束限制了通过多维修队合作减少重要路段修复时间,从而改善修复效果的可能.为解决该问题,基于受损路网模型和多维修队修复同步性,提出多维修队合作修复策略的约束式,研究以路网性能累计效用最大化为目标的应急救援阶段灾后路网多维修队合作修复调度与路由联合优化问题.首先,根据灾后节点可达性建立路网性能评价指标;然后,设计多维修队场景下修复动作的回报函数,以此为基础构建灾后路网多维修队合作修复调度与路由联合优化问题的马尔可夫决策过程,并采用贪心算法和 Q 学习算法求解该问题;最后,通过案例分析结果表明,在随机破坏和灾难点破坏场景中,所提出方法能够加速路网性能的恢复,有效提升应急救援阶段路网性能累计效用,保障灾后救援活动顺利进行.

关键词: 受损路网修复; 联合优化; 多维修队合作; 马尔可夫决策过程; 贪心算法; Q 学习

中图分类号: U418 文献标志码: A

DOI: 10.13195/j.kzyjc.2023.1724

引用格式: 谭晓燕,毛新华,周继彪,等.考虑多维修队合作策略的灾后路网修复调度与路由联合优化[J].控制与决策,2024,39(12):4151-4159.

Joint optimization of scheduling and routing in post-disaster road network repair considering multi-crew cooperation

TAN Xiao-yan^{1,2,3}, MAO Xin-hua^{1,2,3†}, ZHOU Ji-biao⁴, XU Ting¹

- (1. College of Transportation, Chang'an University, Xi'an 710064, China; 2. Engineering Research Center of Digital Construction and Management for Transportation Infrastructure of Shaanxi Province, Xi'an 710064, China; 3. Xi'an Key Laboratory of Digitalization of Transportation Infrastructure Construction and Management, Xi'an 710064, China; 4. Ningbo High Grade Highway Construction Management, Ningbo 315199, China)

Abstract: The common constraint in the research on post-disaster road network repair decision problems is that each damaged road segment is repaired by a single crew, which limits the possibility of reducing the repair time of the important road segment through multi-crew cooperation, so as to improve the repair effect. In order to solve this problem, this paper proposes the constraints of multi-crew cooperation based on the damaged road network model and the synchronization of multi-crew, and studies the multi-crew coordinated scheduling and routing problem in the post-disaster emergency rescue stage with the goal of maximizing the cumulative utility of road network performance. The road network performance index is established according to node accessibility after disaster, and the reward function of repair action under the multi-crew scenario is designed. The Markov decision process of the multi-crew coordinated scheduling and routing problem is constructed, and the greedy algorithm and Q learning algorithm are used to solve the problem. Finally, the case analysis shows that the proposed method can accelerate the repair of road network performance, effectively improve the cumulative utility of road network performance in the emergency rescue stage, and ensure the smooth progress of post-disaster rescue activities, no matter in the random failure scenario or disaster point failure scenario.

Keywords: damaged road network repair; joint optimization; multi-crew cooperation; Markov decision process; greedy algorithm; Q learning

收稿日期: 2023-12-11; 录用日期: 2024-03-11.

基金项目: 国家自然科学基金青年科学基金项目(52102374); 宁波市自然科学基金重点项目(2023J028); 陕西省自然科学研究计划项目(2020JQ-360); 中央高校基本科研业务费专项资金项目(300102343203, 300102343205).

责任编辑: 刘向杰.

†通讯作者. E-mail: maoxinhua@chd.edu.cn.

0 引言

近年来,自然灾害的频发给全世界造成了巨大的人员伤亡和财产损失,其原因往往不是救援物资供应不足,而是路段中断导致救援物资无法运送到受灾需求点^[1],尤其是在欠发达地区和路网密度低、供应链供给不足的农村地区. 尽管灾前防灾减灾措施一定程度上能够提升路网抵抗灾害的能力,但是灾害带来的损害无法完全避免. 基于此,高效且科学地制定应急救援阶段^[2]受损路网的修复策略对快速恢复路网连通,保障灾后救援活动顺利进行至关重要.

灾后路网修复决策方法主要分为修复优先级和优化决策两大类. 前者基于鲁棒性^[3-4]、韧性^[5-8]等指标评估路段重要性,确定受损路段的恢复优先级. 该方法是省时的,但是忽略了受损路段的关系,导致确定的修复策略很少是最优的^[9]. 为了解决上述问题,众学者通过规划模型演化受损路网的修复问题,并设计相应算法求解该问题. 由于该问题有NP-hard难的性质,精确算法求解时间往往以指数级别增长^[1],启发式算法成为求解该问题的常用算法. Li等^[2]和毛新华等^[10-11]将灾后路网修复调度问题表述为以韧性最大化为目标的混合整数规划模型,并设计了遗传算法和禁忌搜索算法求解该问题; Gokalp等^[12]使用动态规划描述灾后路网修复问题,并针对该问题开发了双向搜索启发式算法; Moghtadernejad等^[9]将受损路段分为两个小组,分别求解组内路段修复调度. 实验结果表明,受损路段和维修队数目增加会使得启发式算法求解时间显著增加. 为了解决启发式算法求解速度慢的问题,张国富等^[13-15]探索使用强化学习来求解受损路网修复调度问题,并验证了Q学习和深度Q学习能够在较短时间内求出稳定且可靠的修复策略. 但是该结论仅限于单维修队场景,多维修队场景中强化学习求解效果有待进一步探索.

上述研究聚焦于受损路段的修复调度决策,忽略了修复过程中维修队的路由决策. 由于道路中断,修复次序相邻的受损路段间可能距离较远,甚至无法通行^[16],导致修复调度方案不再最优. Maya-Duque等^[1]定义了灾后路网的单维修队修复调度和路由问题(the single crew scheduling and routing problem, SCSRP),将其表述为动态规划模型,并设计了精确动态规划算法和迭代贪婪随机化构造以求解该问题. 由于SCSRP为多主体优化问题,具有高度复杂性,相关研究使用Benders-cut将SCSRP分解为多个简单问题以便于求解^[17-19]. 虽然SCSRP研究获得了诸多有价值的成果,但是将SCSRP扩展至灾后路网的多维

修队修复调度和路由问题(the multi-crew scheduling and routing problem, MCSRP)的研究是极少的^[20-21]. 另外,现有路网修复决策研究多基于“一个受损路段只能由一个维修队修复”的约束展开,这一约束限制了通过多维修队合作改善受损路网修复效果的可能,有待被松弛^[1,20-21].

为了弥补现有研究的不足,本文研究应急救援阶段灾后路网的多维修队合作修复调度与路由联合优化问题(the multi-crew coordinated scheduling and routing problem, MCCSRP),以实现应急救援阶段路网性能累计效用最大化. 首先,基于灾后需求点可达性建立应急救援阶段的路网性能评价指标,并使用路网性能累积效用评估修复策略的优劣;然后,松弛MCSRP中“一个受损路段只能由一个维修队修复”的约束,并设计多维修队场景下动作回报函数,在此基础上构建MCCSRP的马尔可夫决策过程;最后,设计贪心算法和Q学习算法求解MCCSRP,并在算法中增加非可行动作屏蔽机制以避免算法失效.

1 问题描述

使用无向连通图 $G = (N, A)$ 定义MCCSRP. 其中: N 为节点集合,包括工作站集合、需求点集合和交叉点集合. 工作站为维修人员出发的地点,需求点为灾后需要人道主义救济的地点,交叉点为两个或多个边相交的节点. 一般情况下,需求节点、工作站和交叉点不作区分,因为后两类节点可看作没有需求的需求点^[21]. A 为路段集合,包括受损路段集合和未受损路段集合. 受损路段又分为中断路段集合 A^* 和非中断路段集合. 应急救援阶段的修复活动主要是道路障碍物清理和临时道路搭建,即中断路段的打通,因此 A^* 为待修复路段集合.

将MCCSRP描述为马尔可夫决策过程(Markov decision process, MDP),并定义 $K = \{1, 2, \dots, \kappa\}$ 为MDP的决策点集合. 其中: $\kappa \leq |A^*|$, t_k 为决策点 $k(k \in K)$ 所处时刻, $|A^*|$ 为中断路段总数. t_0 时刻灾害发生,部分道路中断,路网性能 φ 下降至较低水平. t_1 时刻维修队从工作站出发前往中断路段展开修复活动. t_k 时刻维修队 s_w 完成中断路段 a_w 的修复活动,此时,将空闲维修队 s_w 指派至下一个要修复的中断路段 b_w ,并求解两个中断路段间的最短路径 d_w 和最短路径通行时间 γ_w . 其中: $w = 1, 2, \dots, S_k, S_k$ 为 t_k 时刻空闲维修队数目. 重复决策,直至中断路段全部完成修复或修复工期达到应急救援阶段可接受的最长修复工期 T_{\max} . 与资源约束项目调度问题的MDP一致,决策是在每个活动结束时做出的^[22]. 因此, t_k 也是中断路段

a_w 的修复活动完成时刻.

MCCSRP优化决策基于以下假设条件展开:1)中断路段打通后才恢复通行能力;2)维修队同时只能修复一条路段;3)修复活动不能中途终止.

2 模型构建

2.1 目标函数

2.1.1 路网性能

参考文献[23],以中断路段全部打通后节点*i*到其他节点最小路径通行时间之和与*t*时刻节点*i*到其他节点最小路径通行时间之和的比值计算*t*时刻节点*i*的可达率 $\varphi_i(t)$.具体如下所示:

$$\varphi_i(t) = \frac{\sum_{j \in N} D_{ij}}{\sum_{j \in N} D_{ij}(t)}, \forall i \in N. \quad (1)$$

其中: D_{ij} 为中断路段全部打通后节点*i*到节点*j*最小路径通行时间, $D_{ij}(t)$ 为*t*时刻节点*i*到节点*j*最小路径通行时间.当节点*i*与节点*j*不连通时,人员疏散和救援物资运送需要依靠人力,会产生大量的通行时间,因此,在模型中设置一个较大的常数 D_{\max} 来反映节点不连通时的通行时间.

使用节点平均可达率评价路网性能,以反映路网对灾后应急救援活动的支持能力.*t*时刻的路网性能 $\varphi(t)$ 的计算如下所示:

$$\varphi(t) = \frac{\sum_{i \in N} \varphi_i(t)}{|N|}, \quad (2)$$

其中 $|N|$ 为路网中节点总数.

2.1.2 路网性能累积效用

MCCSRP优化决策以应急救援阶段路网性能累积效用 R^* 最大化为目标.基于假设1),修复过程中路网性能呈阶梯式增长,将 R^* 表示为

$$R^* = \sum_{k \in K} \varphi(t_k) \times (t_{k+1} - t_k). \quad (3)$$

为了使得不同灾害场景下应急救援阶段路网性能累积效用具有可比性,将 R^* 归一化处理得到

$$R = \frac{\sum_{k \in K} \varphi(t_k) \times (t_{k+1} - t_k)}{\varphi_0 \times T_{\max}}. \quad (4)$$

其中: R 为一个无量纲的值, R 越接近1,应急救援阶段路网性能累积效用越大,节点可访问性也越好; φ_0 为中断路段全部打通后的路网性能;当 $k = \kappa$ 时,定义 $t_{k+1} = t_1 + T_{\max}$.

2.2 马尔可夫决策过程

将MDP表示为一个四元组 $\langle C, O, \pi, r \rangle$.其中: C 和 O 分别为状态空间和可行动作空间, π 为状态转移概率, r 为奖励函数.

2.2.1 状态

设计 $C_k = \langle t_k, F_k, P_k \rangle$ 描述系统在 t_k 时刻的状态.变量 $f_s^k \in F_k (\forall s \in S)$ 描述 t_k 时刻维修队*s*的位置, S 为维修队集合.变量 $p_a^k \in P_k (\forall a \in A^*)$ 描述 t_k 时刻中断路段*a*的状态,具体定义如下所示:

$$p_a^k = \begin{cases} -2, & a \text{在} t_k \text{时刻前被打通;} \\ -1, & a \text{在} t_k \text{时刻被打通;} \\ 0, & a \text{未被指派维修队;} \\ H, & a \text{在} H \text{时刻被打通, } H > t_k; \end{cases} \quad \forall a \in A^*. \quad (5)$$

2.2.2 可行动作

给定状态 C_k 下,需要为空闲维修队确定下一个要维修的中断路段和最短路径,因此, $\{b_1, b_2, \dots, b_{S_k}\}$ 和 $\{d_1, d_2, \dots, d_{S_k}\}$ 为状态 C_k 下的可行动作,记为 o_k ($o_k \in O$).为了简化表述,定义 $B_k = \{b_1, b_2, \dots, b_{S_k}\}$.

MCSR P中“一个受损路段只能由一个维修队修复”的具体约束式如下所示:

$$p_a^k = 0, \forall a \in B_k; \quad (6)$$

$$U_a^k \leq 1, \forall a \in B_k. \quad (7)$$

其中:约束(6)表示维修队不能去往 t_k 时刻正在被修复或已完成修复的中断路段,约束(7)表示 t_k 时刻指派给中断路段*a*的维修队数目 U_a^k 不超过1个.

MCCSRP将“一个受损路段只能由一个维修队修复”松弛为“维修队到达待修复路段时,该路段修复活动未结束”,通过扩大可行域求得更好的优化解,实现路网修复效果的提升.松弛后约束式如下所示:

$$p_a^k \geq 0, \forall a \in B_k. \quad (8)$$

$$U_a^k = \sum_{a^k \in B_k} [a^k = a], \forall a \in B_k. \quad (9)$$

$$V_a^k = \sum_{s \in S} [f_s^k = a], \forall a \in B_k. \quad (10)$$

$$p_a^k > t_k + d_{a,1}^k, \forall \{a \in B_k | p_a^k > 0\}. \quad (11)$$

$$M_a^{u-1} > t_k + d_{a,u}^k, \forall \{a \in B_k | U_a^k \geq 2\}, \\ \{u \in Z^+ | 2 \leq u \leq U_a^k\}. \quad (12)$$

$$M_a^1 = \begin{cases} t_k + d_{a,1}^k + \omega_a, & p_a^k = 0; \\ p_a^k - \frac{p_a^k - (t_k + d_{a,1}^k)}{V_a^k + 1}, & p_a^k > 0; \end{cases} \quad \forall a \in B_k. \quad (13)$$

$$M_a^u = M_a^{u-1} - \frac{M_a^{u-1} - (t_k + d_{a,u}^k)}{u + V_a^k},$$

$$\forall \{a \in B_k | U_a^k \geq 2\}, \{u \in Z^+ | 2 \leq u \leq U_a^k\}. \quad (14)$$

$$V_a^k + U_a^k \leq S_{\max}^a, \forall a \in B_k. \quad (15)$$

其中:约束(8)表示不能指派维修队去往 t_k 时刻已完成修复的中断路段.式(9)计算 t_k 时刻指派给中断路

段 a 的维修队数目 U_a^k . 这里: $[\cdot]$ 为一个判断变量, 若 $[\cdot]$ 内式子成立, 则 $[\cdot] = 1$; 否则, $[\cdot] = 0$. 约束 (10) 表示计算 t_k 时刻正在修复中断路段 a 的维修队数目 V_a^k . 约束 (11) 和 (12) 保证新增维修队到达正在被修复的中断路段 a 时, 中断路段 a 的修复活动未结束. 这里: M_a^u 为 $u + V_a^k$ 个维修队合作修复中断路段 a 时, 中断路段 a 的修复完成时刻. 使用一组有规律的计算式 (13) 和 (14) 求解, 通过在 M_a^{u-1} 的基础上减去增加 1 个维修队后所减少的修复时间而得到, 增加的是 t_k 时刻指派给中断路段 a 的维修队中第 u 个到达路段 a 的维修队, 该维修队从上一个中断路段到中断路段 a 最小路径通行时间记为 $d_{a,u}^k$. 另外, 式 (13) 中, ω_a 为中断路段 a 仅被一个维修队修复所需要的时间. 约束 (15) 表示同时修复中断路段 a 的维修队数目不能超过该路段允许的最大数目 S_{\max}^a , 反映较窄路段或灾后结构不稳定路段不能被过多维修队同时修复的实际情况.

修复动作需满足修复工期以及可达性约束为

$$M_a^{U_a^k} \leq t_1 + T_{\max}, \forall a \in B_k; \quad (16)$$

$$\gamma < M, \forall \gamma \in \{\gamma_1, \gamma_2, \dots, \gamma_{S_k}\}. \quad (17)$$

其中: 约束 (16) 规定整个修复工期不超过 T_{\max} , 约束 (17) 保证维修队只在两个连通的中断路段间转移, M 为一个足够大的数.

最短路径 $\{d_1, d_2, \dots, d_{S_k}\}$ 和最短路径通行时间 $\{\gamma_1, \gamma_2, \dots, \gamma_{S_k}\}$ 由如下最短路径模型确定:

$$\min \gamma_l = \sum_{a_{ij} \in A} Y_{a_{ij}}^{s_l} Z_{a_{ij}}^k. \quad (18)$$

$$\text{s.t. } Y_{a_{ij}}^{s_l} = \{0, 1\}, \forall a_{ij} \in A. \quad (19)$$

$$\sum_{a_{ij} \in A} Y_{a_{ij}}^{s_l} - \sum_{a_{ji} \in A} Y_{a_{ji}}^{s_l} = \begin{cases} 1, & i = n_{a_l}; \\ -1, & i = n_{b_l}; \\ 0, & \text{otherwise}; \end{cases} \forall i \in N. \quad (20)$$

$$\sum_{a_{ij} \in A} Y_{a_{ij}}^{s_l} + \sum_{a_{ji} \in A} Y_{a_{ji}}^{s_l} \leq 2, \forall i \in N. \quad (21)$$

$$Z_{a_{ij}}^k = \begin{cases} Z_{a_{ij}}, & p_{a_{ij}} < 0; \\ M, & \text{otherwise}; \end{cases} \forall a_{ij} \in A^*. \quad (22)$$

$$Z_{a_{ij}}^k = Z_{a_{ij}}, \forall a_{ij} \in A \setminus A^*. \quad (23)$$

$$Z_{a_l}^k = Z_{b_l}^k = 0. \quad (24)$$

其中: 目标函数 (18) 最小化状态 C_k 下, 空闲维修队 s_l 从中断路段 a_l 转移至 b_l 的通行时间. 这里: a_{ij} 表示以点 i 为起点, 点 j 为终点的路段; $Z_{a_{ij}}^k$ 为状态 C_k 下路段 a_{ij} 的通行时间. 约束 (19) 保证 $Y_{a_{ij}}^{s_l}$ 为二进制决策变量, 若维修队 s_l 从路段 a_l 转移至 b_l 的路径中包含路段 a_{ij} , 则 $Y_{a_{ij}}^{s_l} = 1$; 否则, $Y_{a_{ij}}^{s_l} = 0$. 约束 (20) 表示流平衡.

这里: n_{a_l} 为 a_l 的任意一个端点, n_{b_l} 为 b_l 的任意一个端点. 约束 (21) 规定每个路段最多被访问 1 次. 式 (22) 和 (23) 计算 t_k 时刻路段阻抗, 这里 $Z_{a_{ij}}$ 为路段未中断时的通行时间. 式 (24) 将起点路段和终点路段的通行时间设置为 0, 使得维修队 s_l 可在 a_l 任意一端出发, 也可到达 b_l 任意一端.

2.2.3 状态转移概率

MDP 中 t_{k+1} 时刻的状态 C_{k+1} 由 t_k 时刻的状态 C_k 和所选择的动作 o_k 唯一确定. 状态转移方程记为 $C_{k+1} = \Pi(C_k, o_k)$, 具体的对应关系如下所示:

$$t_{k+1}^1 = \min_{\{p_a^k \in P_k | p_a^k > 0\}} p_a^k. \quad (25)$$

$$t_{k+1}^2 = \min_{a \in B_k} M_a^{U_a^k}. \quad (26)$$

$$t_{k+1} = \min\{t_{k+1}^1, t_{k+1}^2\}. \quad (27)$$

$$f_{s_l}^{k+1} = b_l, \forall 1 \leq l \leq S_k. \quad (28)$$

$$f_{s_l}^{k+1} = f_{s_l}^k, \forall s \in S \setminus \{s_1, s_2, \dots, s_{S_k}\}. \quad (29)$$

$$p_a^{k+1} = \begin{cases} -1, & M_a^{U_a^k} = t_{k+1}; \\ M_a^{U_a^k}, & \text{otherwise}; \end{cases} \forall a \in B_k. \quad (30)$$

$$p_a^{k+1} = \begin{cases} -2, & p_a^k = -1; \\ -1, & p_a^k = t_{k+1}; \\ p_a^k, & \text{otherwise}; \end{cases} \forall a \in A^* \setminus B_k. \quad (31)$$

其中: 式 (25)~(27) 计算决策点 $k+1$ 的决策时刻, 等于下一个中断路段修复活动的结束时刻; 式 (28) 和 (29) 描述维修队的位置变化; 式 (30) 描述 t_{k+1} 时刻中断路段 a ($\forall a \in B_k$) 的状态变化; 式 (31) 描述中断路段 a ($\forall a \in A^* \setminus B_k$) 的状态, 这里修复结束时刻等于 t_k 或 t_{k+1} 的中断路段状态发生改变.

转移概率 $\pi_{cc'}^o$ 为在决策点 k 执行动作 o 后, 在 t_{k+1} 时刻系统状态从 c 转移至 c' 的概率. 基于状态转移方程, 得到

$$\pi_{cc'}^o = \begin{cases} 1, & c' = \Pi(c, o); \\ 0, & \text{otherwise}; \end{cases} \forall c' \in C. \quad (32)$$

值得注意的是, MCCSRP 中的系统状态数目众多, 且很难提前全部演化出来, 这使得决策者难以事先获取到状态转移概率矩阵.

2.2.4 动作回报

基于决策目标, 设计路网性能累计效用增加值来衡量状态 C_k 下执行动作 o_k 的瞬时回报 $r(C_k, o_k)$, 有

$$r(C_k, o_k) = R_{k, o_k}^* - R_k^*. \quad (33)$$

其中: $\varphi_0 \times T_{\max} \geq r(C_k, o_k) \geq 0$, 且 $r(C_k, o_k)$ 越大, 状态 C_k 下执行动作 o_k 的瞬时回报越高; R_{k, o_k}^* 为 t_k 时刻执行动作 o_k 后路网性能累计效用; R_k^* 为 t_k 时刻不执行任何动作下路网性能累计效用. 使用两个维修队的

简单例子具体说明 R_{k,o_k}^* 和 R_k^* 的计算. 路网性能累计效用变化如图1所示. 若维修队在 t_1 时刻不执行任何动作, 则路网性能将维持在 $\varphi(t_1)$, 路网性能累计效用为竖线填充区域面积 ψ_1 , 即 $R_1^* = \psi_1$; 若执行动作 o_1 , 分别在 t_2 和 t_{12} 时刻完成修复活动, 则路网性能累计效用增长为竖线填充区域面积和点线填充区域面积 ψ_2 之和, 即 $R_{1,o_1}^* = \psi_1 + \psi_2$.

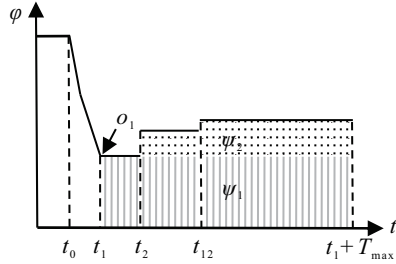


图1 路网性能累计效用变化

3 算法

3.1 非可行动作屏蔽机制

为了避免不可行动作造成算法停滞或失效, 本文将MDP的动作空间定义为可行动作空间. 遍历所有动作来确定可行动作空间是有效的, 但是MCCSRP的动作空间较大, 使得遍历法较为耗时. 为此, 在贪心算法和Q学习算法中增加非可行动作屏蔽机制, 以提高可行动作空间确定效率. 首先, 屏蔽 C_k 状态已完成修复的中断路段, 以及合作维修队数目到临界值的中断路段, 确保动作满足约束(8)和(15); 然后, 使用Dijkstra's算法^[25]计算空闲维修队转移至中断路段的最短路径以及最短路径通行时间, 屏蔽不可达的中断路段, 确保动作满足约束(17); 最后, 将每个空闲维修队可达的中断路段交叉组合得到动作集合, 并屏蔽不满足合作约束(9)~(14)和修复工期约束(16)的动作, 得到 C_k 状态下可行动作空间 O_{C_k} .

3.2 贪心算法

为了满足应急救援阶段灾后路网修复决策对低时间复杂度和低空间复杂度的要求, 本节基于动作回报提出一种贪心算法. 考虑修复过程中状态的变化, 在修复活动完成时计算该状态下可行动作的奖励值, 应用贪心原则选取当前状态下最优动作, 快速且高效地确定MCCSRP的近似最优策略, 提高应急救援阶段路网性能累计效用. 具体求解步骤如下.

step 1: 输入受损路网模型, 并初始化状态 $C = C_1(t_1 = 0; f_s^1 = \text{工作站}, \forall s \in S; p_a^1 = 0, \forall a \in A^*)$.

step 2: 创建最优动作集合 $X = \emptyset$, 最优动作回报集合 $Y = \emptyset$.

step 3: 计算 O_C , 若 $O_C = \emptyset$, 则转至 step 5; 否则, 计算动作奖励值 $r(C, o) (\forall o \in O_C)$.

step 4: 执行动作 o^* , $r(C, o^*) = \max_{o \in O_C} r(C, o)$, 更新集合 $X = X \cup o^*$, $Y = Y \cup r(C, o^*)$.

step 5: 观察下一状态 C' .

step 6: 判断是否满足决策终止条件(中断路段全部完成修复或修复工期达到 T_{\max}), 若满足, 则输出 X 和 Y ; 否则, 更新 $C = C'$, 转至 step 3.

3.3 Q学习算法

基于MCCSRP的MDP模型中状态转移概率矩阵难以事先获取的特点, 使用无模型强化学习中的Q学习算法^[26]求解该问题. Q学习算法的主要思想是建立一个状态和动作组成的Q表, 通过智能体与环境交互更新Q表, 实现对Q值的估计, 获得最优策略. 基于贝尔曼方程的Q值迭代公式如下所示:

$$Q(C_k, o_k) = (1 - \alpha)Q(C_k, o_k) + \alpha(r(C_k, o_k) + \gamma \max_{o \in O_{C_{k+1}}} Q(C_{k+1}, o)). \quad (34)$$

其中: $Q(C_k, o_k)$ 为 C_k 状态下执行动作 o_k 获得收益的期望值; $0 \leq \alpha \leq 1$ 为学习速率, 用于调控学习前期产生的奖励对当前Q值更新的影响; γ 为折扣因子, 表示未来回报与当前回报间的差异.

设置一个经验池用于存储智能体学习过程中探索到的状态以及对应的可行动作^[16]. 当智能体探索到新状态时, 计算该状态的可行动作空间, 并存储至经验池中, 然后在Q表中存储该状态可行动作的初始Q值. 该方法使得Q学习能够适应MCCSRP中状态以及可行动作空间不能提前演化出来的特点, 还避免了智能体学习过程中重复为同一状态确定可行动作空间带来的时间损耗. 改进Q学习算法具体步骤如下.

step 1: 输入受损路网模型参数, 并初始化经验池 Ω 和Q表, 学习次数 $\text{epoch} = 1$.

step 2: 设置初始状态 $C = C_1$.

step 3: 若 $C \notin \Omega$, 则计算 O_C , 在 Ω 中存储 C 和 O_C , 初始化对应Q值为0; 否则, 跳过该步骤.

step 4: 若 $O_C = \emptyset$, 则跳过该步骤; 否则, 按 ε -贪心策略^[27]选择动作.

step 5: 观察下一状态 C' , 并由式(33)计算动作回报, 由式(34)更新Q值.

step 6: 判断是否满足决策终止条件, 若满足, 则计算 epoch 轮学习到修复策略的目标函数值, 以及Q表对应最优修复策略的目标函数值; 否则, 更新状态 $C = C'$, 转至 step 3.

step 7: $\text{epoch} = \text{epoch} + 1$, 若达到最大学习周期, 则智能体学习结束, 输出Q表和 Ω ; 否则, 转至 step 2.

step 8: 输出Q表对应最优修复策略^[16].

4 案例研究

4.1 灾害场景

在Sioux Fall网络^[28]中进行案例研究. 使用灾难点攻击和随机攻击两种方式模拟如表1所示的6个受损路网场景, 其中灾难点攻击旨在模拟滑坡、地震等对路网造成聚集性破坏的灾害. 首先, 在路网中随机选择一个节点代表灾难中心; 然后, 依路段中断的概率选取部分路段进行中断破坏, 路段中断的概率与其到灾难中心最小路径通行时间成反比^[12]. 随机攻击是在路网中随机选取部分路段进行中断破坏, 该方法能够提供中断路段相对稀疏的破坏场景, 有助于模拟灾害对路网影响的随机性.

表1 灾害场景

场景	攻击方式	中断路段比例/%	$\varphi(t_1)$	工作站	T_{\max}/h
1	灾难点攻击	20	0.411 2	10	24
2	随机攻击	20	0.422 8	10	24
3	灾难点攻击	30	0.388 2	20	24
4	随机攻击	30	0.257 1	10	24
5	灾难点攻击	40	0.091 5	20	48
6	随机攻击	40	0.071 0	20	48

考虑到未中断路段虽然可以通行, 但是通行速度和能力往往均会减少, 假设未中断路段的通行能力和通行速度减为灾前的一半, 且中断路段打通后其通行能力和通行速度也为灾前的一半. 中断路段允许合作维修队的最大数目 S_{\max}^a 由灾前路段的容量确定, 假定容量偏大的半数中断路段 $S_{\max}^a = 2$, 容量偏小的半数中断路段 $S_{\max}^a = 1$. 另外, 使用灾后路网节点的接近中心性评价节点中转功能, 并选取接近中心性最大的节点为工作站. 假定有两个维修队在 $t_1 = 0$ 时刻开始应急救援阶段的路网修复活动.

将中断路段修复工期定义为全程由一个维修队修复所需的时间. 为确保合理性, 修复工期根据每个路段通行能力和通行时间的乘积来确定, 具有较高乘积中断路段的修复工期在较高均值和方差的伽马分布中采样, 而较低乘积中断路段的修复工期则从具有较低均值和方差的伽马分布中采样^[12]. 这反映出容量越大、线路越长, 越需要更多的劳动力来重建. 应急救援阶段路网修复工作以道路障碍物清理和临时道路搭建为主, 因此, 将修复工期的单位设置为小时, 且对应修改文献[12]中伽马分布参数的单位, 进而得到本文受灾场景下修复工期的采样参数.

4.2 实验结果与分析

进行两组实验来验证所提出方法的有效性. 实验1分析修复优先级、贪心算法、 Q 学习算法求解结果的差异. 实验2评价MCSR和MCCSR最优修复

策略的修复效果. 实验代码在Matlab R 2023 a中编码并执行. 实验环境为Intel(R) Core(TM) i7-7770 CPU 3.6 GHz, 内存16 GB的Windows操作系统. 模型参数为 $D_{\max} = 10$. Q 学习算法参数如下: $\varepsilon = 0.3$, $\alpha = 0.7$, $\gamma = 0.95$.

4.2.1 实验1

使用修复优先级、贪心算法和 Q 学习算法分别求解6个灾害场景的最优修复策略, 其中修复优先级使用打通中断路段后 R_1^* 的提升值评价路段重要性^[7]. 3种算法求得最优修复策略的目标函数值如表2所示. 贪心算法和 Q 学习算法求得的最优修复策略均优于修复优先级求得的最优策略. 另外, 相较于贪心算法, Q 学习算法在6种场景下均获得了更好的结果, 目标函数值的提升率在0.19%~4.49%间.

表2 不同算法最优策略的目标函数值

场景	修复优先级	贪心算法	Q 学习算法	提升率/%
1	0.721 6	0.759 5	0.792 2	4.30
2	0.811 1	0.827 4	0.831 3	0.47
3	0.696 5	0.771 0	0.798 8	3.60
4	0.589 8	0.652 0	0.654 6	0.40
5	0.461 8	0.762 4	0.763 9	0.19
6	0.550 9	0.656 3	0.685 8	4.49

图2为3种算法求得最优修复策略下路网性能的变化情况. 修复开始时: 3种算法下路网性能增长趋势相近, 但是随着修复的进行, 贪心算法和 Q 学习算法对应最优修复策略下路网性能相对更加快速地恢复, 且恢复至更高的水平. 这一现象反映出, 随着系统状态的改变, 路段关键程度会发生变化, 导致修复优先级方法确定的修复策略不是最优的, 甚至不是近似最优. 贪心算法对应路网修复曲线与 Q 学习算法对应修复曲线的差异性较小, 表明贪心算法考虑状态变化动态评价动作的重要性一定程度上能够解决修复优先级方法的缺陷, 获取到近似最优解.

为了深入分析 Q 学习算法在多维修队场景下的求解效率, 图3展示了不同场景下智能体的学习过程. 智能体学习前期获得的修复策略中部分优于对应 Q 表中的最优策略, 一定周期数时, 智能体不会学习到比 Q 表中的最优策略更优的策略, Q 表对应的最优策略也不再改变, 此时, 得到最优修复策略. 在中断率小于30%的场景中, 智能体在1000个学习周期内学习到最优修复策略, 但是在中断率为40%的场景需要3000个左右学习周期学习到最优修复策略. 这是因为多维修队场景下待修复路段的增加使得智能体的状态集与动作集规模呈指数增长, 智能体需要更长时间学习.

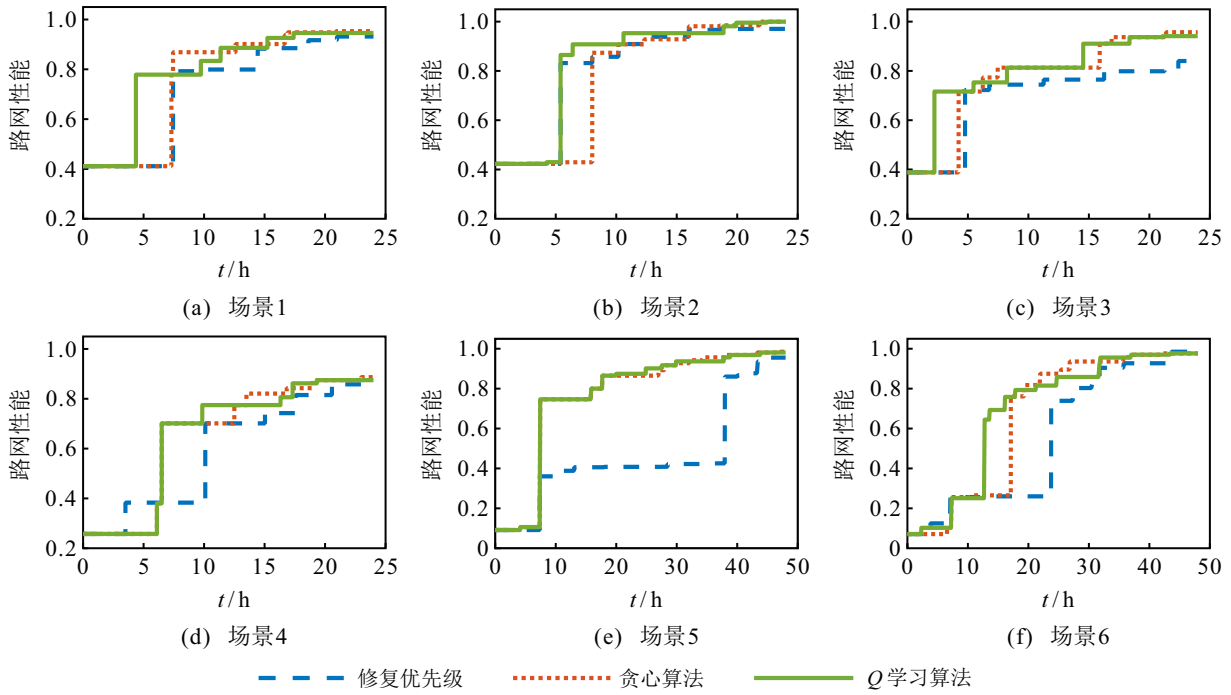


图2 不同算法下路网性能的变化

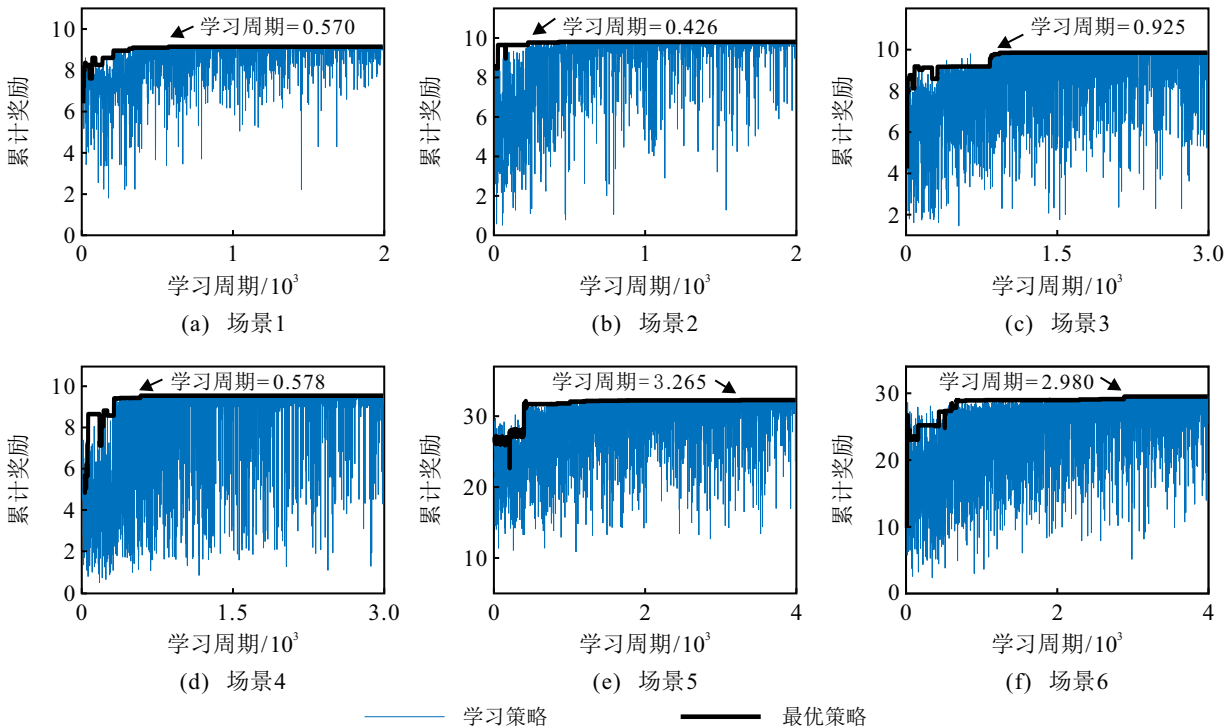


图3 不同场景智能体的学习过程

4.2.2 实验2

为了验证多维修队合作修复策略的必要性和有效性,分别在6个受损场景中使用Q学习算法求解MCSR和MCCSR.表3为MCSR和MCCSR最优策略目标值.由表3可见:MCCSR得到的最优修复策略目标函数值均大于等于MCSR,表明在灾后路网修复调度和路由决策优化问题中考虑多维修队合作策略是有效且必要的.

图4为两种方法对应最优修复策略实施下路

网性能的变化情况,以直观的形象展示MCCSR和MCSR的修复效果.由图4可见:MCCSR对应路网

表3 MCSR和MCCSR最优策略目标值

场景	MCSR	MCCSR	提升率/%
1	0.7596	0.7922	4.28
2	0.8270	0.8313	0.52
3	0.7710	0.7988	3.61
4	0.6521	0.6546	0.39
5	0.7607	0.7639	0.42
6	0.6563	0.6858	4.49

性能曲线较MCSR P对应路网性能曲线向左偏,反映出MCCSR P通过调用多个维修队合作修复同一个中断路段,缩短了重要路段修复工期,尽早实现了路网性能的增加.另外,MCCSR P较MCSR P的优势在

场景1、场景3和场景6中更为突出,表明相较于修复后期,修复前期多维修队合作对路网性能累计效用的提升效果更为显著,这是由于修复后期路网性能的增加速度相较于修复前期更为缓慢.

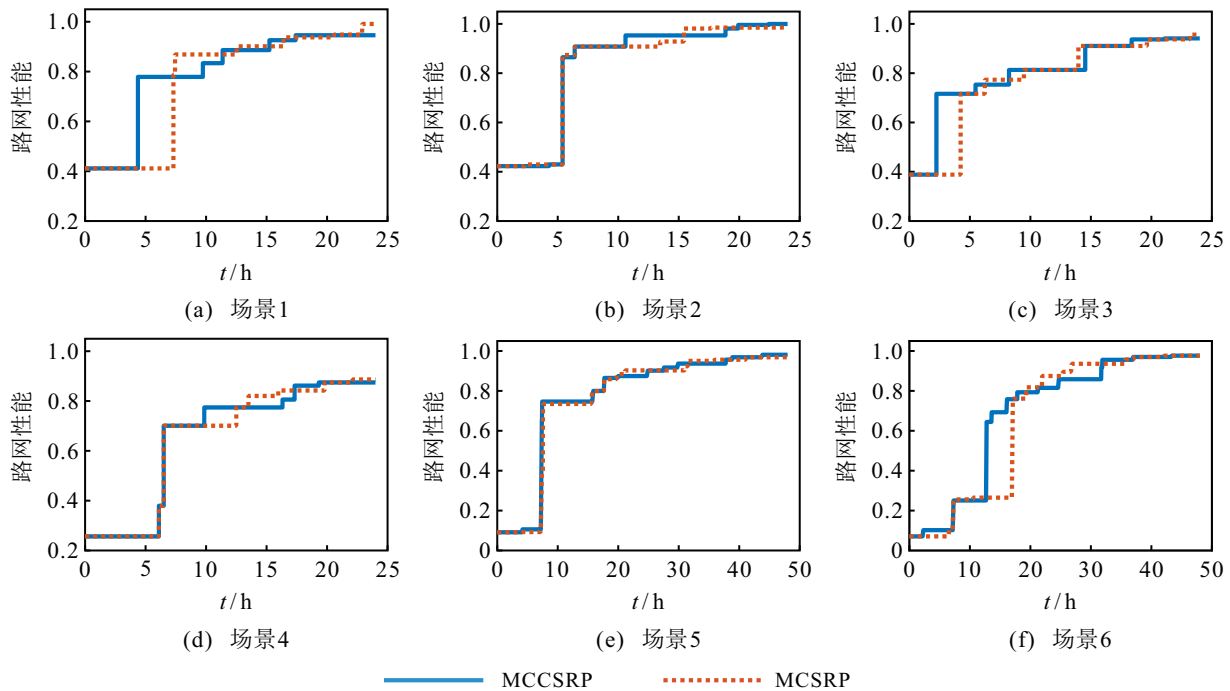


图4 MCCSR P和MCSR P下路网性能的变化

5 结论

本文以路网性能累计效用最大化为目标,研究了应急救援阶段的MCCSR P.使用MDP演化MCCSR P的决策过程,并设计了贪心算法和 Q 学习算法进行求解.在随机破坏场景和灾难点破坏场景中的实验结果表明,所构建的MCCSR P通过减少重要路段的修复时间来加速路网性能的恢复,可有效提升应急救援阶段路网性能累计效用.另外,贪心算法和 Q 学习算法均展示出较好的稳定性和可靠性,且 Q 学习算法获取到的最优修复策略优于贪心算法获取到的最优修复策略.但是随着待修复路段的增加,智能体需花费更长时间来学习,使得 Q 学习算法相对更适合在小型受灾场景或决策时间充足的情况下应用,时间复杂度和空间复杂度双低的贪心算法更适合在大型受灾场景或决策时间不充足的情况下应用.本文对强化学习在多维修队路网修复决策问题中的应用仅做了初步探索,如何提升智能体训练效率,以高效求解具有大规模状态集和动作集的路网修复决策问题,将是下一步需要探讨的难题.

参考文献(References)

[1] Maya-Duque P, Dolinskaya I S, Sørensen K. Network repair crew scheduling and routing for emergency relief distribution problem[J]. *European Journal of Operational*

Research, 2016, 248(1): 272-285.
 [2] Li Z L, Jin C, Hu P, et al. Resilience-based transportation network recovery strategy during emergency recovery phase under uncertainty[J]. *Reliability Engineering & System Safety*, 2019, 188: 503-514.
 [3] Scott D M, Novak D C, Aultman-Hall L, et al. Network robustness index: A new method for identifying critical links and evaluating the performance of transportation networks[J]. *Journal of Transport Geography*, 2006, 14(3): 215-227.
 [4] Du M Q, Jiang X W, Cheng L. Alternative network robustness measure using system-wide transportation capacity for identifying critical links in road networks[J]. *Advances in Mechanical Engineering*, 2017, 9(4): 168781401769665.
 [5] Wu Y Y, Hou G Y, Chen S R. Post-earthquake resilience assessment and long-term restoration prioritization of transportation network[J]. *Reliability Engineering & System Safety*, 2021, 211: 107612.
 [6] Barker K, Ramirez-Marquez J E, Rocco C M. Resilience-based network component importance measures[J]. *Reliability Engineering & System Safety*, 2013, 117: 89-97.
 [7] Liu Y C, McNeil S, Hackl J, et al. Prioritizing transportation network recovery using a resilience measure[J]. *Sustainable and Resilient Infrastructure*, 2022, 7(1): 70-81.
 [8] Xu Z P, Ramirez-Marquez J E, Liu Y, et al. A new resilience-based component importance measure

- for multi-state networks[J]. *Reliability Engineering & System Safety*, 2020, 193: 106591.
- [9] Moghtadernejad S, Adey B T, Hackl J. Determination of postdisaster restoration programs for road networks using a double-stage optimization approach[J]. *Journal of Infrastructure Systems*, 2022, 28(3): 04022025.
- [10] 毛新华, 王建伟, 袁长伟, 等. 基于韧性最优的灾后公路网修复调度研究[J]. *中国公路学报*, 2022, 35(6): 289-298.
(Mao X H, Wang J W, Yuan C W, et al. Restoration scheduling for post-disaster road networks based on resilience optimization[J]. *China Journal of Highway and Transport*, 2022, 35(6): 289-298.)
- [11] Mao X H, Zhou J B, Yuan C W, et al. Resilience-based optimization of postdisaster restoration strategy for road networks[J]. *Journal of Advanced Transportation*, 2021, 2021: 1-15.
- [12] Gokalp C, Patil P N, Boyles S D. Post-disaster recovery sequencing strategy for road networks[J]. *Transportation Research—Part B: Methodological*, 2021, 153: 228-245.
- [13] 张国富, 涂冰花, 苏兆品, 等. 一种面向严重受损路网的抢修队调度算法[J]. *控制与决策*, 2021, 36(7): 1663-1671.
(Zhang G F, Tu B H, Su Z P, et al. An algorithm for repair crew scheduling on severely damaged road network[J]. *Control and Decision*, 2021, 36(7): 1663-1671.)
- [14] 张国富, 沈宇锋, 宋晓晓, 等. 受灾路网抢修队动态调度问题的建模与求解方法[J]. *计算机工程*, 2023, 49(6): 300-313.
(Zhang G F, Shen Y F, Song X X, et al. Method for modeling and solving the dynamic scheduling problem of the repair crew for restoring damaged road network[J]. *Computer Engineering*, 2023, 49(6): 300-313.)
- [15] 张国富, 常加远, 苏兆品, 等. 大量需求点下基于深度Q学习的受损路网抢修队调度[J]. *控制与决策*, 2022, 37(12): 3267-3277.
(Zhang G F, Chang J Y, Su Z P, et al. Repair crew scheduling for damaged road network with enormous demand points using deep Q-learning[J]. *Control and Decision*, 2022, 37(12): 3267-3277.)
- [16] 郝新军, 毛新华, 谭晓燕, 等. 基于强化学习的灾后公路网修复决策研究[J]. *中国公路学报*, 2023, 36(8): 292-304.
(Hao X J, Mao X H, Tan X Y, et al. Post-disaster highway network restoration decision based on reinforcement learning[J]. *China Journal of Highway and Transport*, 2023, 36(8): 292-304.)
- [17] Moreno A, Munari P, Alem D. Decomposition-based algorithms for the crew scheduling and routing problem in road restoration[J]. *Computers & Operations Research*, 2020, 119: 104935.
- [18] Moreno A, Munari P, Alem D. A branch-and-benders-cut algorithm for the crew scheduling and routing problem in road restoration[J]. *European Journal of Operational Research*, 2019, 275(1): 16-34.
- [19] Lakzaei S, Rahmani D, Tosarkani B M, et al. Integrated optimal scheduling and routing of repair crew and relief vehicles after disaster: A novel hybrid solution approach[J]. *Annals of Operations Research*, 2023, 328(2): 1495-1522.
- [20] Moreno A, Alem D, Gendreau M, et al. The heterogeneous multicrew scheduling and routing problem in road restoration[J]. *Transportation Research—Part B: Methodological*, 2020, 141: 24-58.
- [21] Li S L, Ma Z J, Teo K. A new model for road network repair after natural disasters: Integrating logistics support scheduling with repair crew scheduling and routing activities[J]. *Computers & Industrial Engineering*, 2020, 145: 106506.
- [22] Peng W L, Lin X J, Li H T. Critical chain based proactive-reactive scheduling for resource-constrained project scheduling under uncertainty[J]. *Expert Systems with Applications*, 2023, 214: 119188.
- [23] Chang S E, Nojima N. Measuring post-disaster transportation system performance: The 1995 Kobe earthquake in comparative perspective[J]. *Transportation Research—Part A: Policy and Practice*, 2001, 35(6): 475-494.
- [24] Özdamar L, Tüzün A D, Ergüneş B. Coordinating debris cleanup operations in post disaster road networks[J]. *Socio-Economic Planning Sciences*, 2014, 48(4): 249-262.
- [25] Dijkstra E W. A note on two problems in connexion with graphs[J]. *Numerische Mathematik*, 1959, 1(1): 269-271.
- [26] Clifton J, Laber E. Q-learning: Theory and applications[J]. *Annual Review of Statistics and Its Application*, 2020, 7: 279-301.
- [27] 温广辉, 杨涛, 周佳玲, 等. 强化学习与自适应动态规划: 从基础理论到多智能体系统中的应用进展综述[J]. *控制与决策*, 2023, 38(5): 1200-1230.
(Wen G H, Yang T, Zhou J L, et al. Reinforcement learning and adaptive/approximate dynamic programming: A survey from theory to applications in multi-agent systems[J]. *Control and Decision*, 2023, 38(5): 1200-1230.)
- [28] Transportation Networks for Research Core Team. Transportation networks for research[EB/OL]. (2021-10-22)[2023-12-01]. <https://github.com/bstabler/TransportationNetworks>.

作者简介

谭晓燕(1999—), 女, 博士生, 主要研究方向为交通基础设施数字化、交通韧性优化, E-mail: 2021134015@chd.edu.cn;

毛新华(1986—), 男, 副教授, 博士, 主要研究方向为交通基础设施数字化、交通韧性优化, E-mail: maoxinhua@chd.edu.cn;

周继彪(1986—), 男, 副教授, 博士, 主要研究方向为交通运输规划与管理、行人交通仿真, E-mail: zhoubiao@tongji.edu.cn;

徐婷(1983—), 女, 教授, 博士, 主要研究方向为交通安全领域、路网优化, E-mail: xuting@chd.edu.cn.